MEME16203Linear Models Marking Guide

Assignment 3

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Session:	May 2023
Due by:	28/7/2023		

Q1. Consider the following models:

Model A:	Model B:
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} = \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 0 \\ 1 & -2 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\epsilon}$	$\mathbf{Y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon} = \begin{bmatrix} 1&1&0&0&1&0&0 \\ 1&1&0&0&0&1&0 \\ 1&1&0&0&0&0&1 \\ 1&0&1&0&1&0&0 \\ 1&0&1&0&0&1&0 \\ 1&0&1&0&0&0&1 \\ 1&0&0&1&1&0&0 \\ 1&0&0&1&0&1&0 \\ 1&0&0&1&0&0&1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} + \boldsymbol{\epsilon}.$

where $\boldsymbol{\epsilon} \sim N(0, \sigma^2 \mathbf{I})$

(a) Find matrix $\mathbf{G}$ such that $\mathbf{X} = \mathbf{WG}$. (10 marks)

Ans.

$$\mathbf{X} = \mathbf{W} \begin{bmatrix} 1 & 1 & 0 \\ 0 & -3 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \mathbf{WG}$$

(b) What is the distribution of $\dfrac{\mathbf{y}^{\mathbf{T}}(\mathbf{I}-\mathbf{P_X})\mathbf{P_W}(\mathbf{I}-\mathbf{P_X})\mathbf{y}}{\sigma^2}$? (10 marks)

Ans.

In this case,

$\mathbf{A} = \dfrac{(\mathbf{I}-\mathbf{P_X})\mathbf{P_W}(\mathbf{I}-\mathbf{P_X})}{\sigma^2}$ and $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$.

$\mathbf{A}\boldsymbol{\Sigma} = (\mathbf{I}-\mathbf{P_X})\mathbf{P_W}(\mathbf{I}-\mathbf{PX}) = (\mathbf{I}-\mathbf{P_X})(\mathbf{P_W}-\mathbf{P_W}\mathbf{P_X})$

Note that since $\mathbf{X} = \mathbf{WF}$, then

$\mathbf{P_W X} = \mathbf{P_W W F} = \mathbf{W F} = \mathbf{X}$ and hence
$\mathbf{P_W P_X} = \mathbf{P_W X (X^T X)^{-1} X^T} = \mathbf{X (X^T X)^{-1} X^T} = \mathbf{P_X}$
$\mathbf{P_X P_W} = \mathbf{P_X^T P_W^T} = \mathbf{(P_W P_X)^T} = \mathbf{P_X^T} = \mathbf{P_X}$

$$\begin{aligned}\mathbf{A\Sigma} &= \mathbf{(I - P_X) P_W (I - PX)}\\ &= \mathbf{(I - P_X)(P_W - P_W P_X)}\\ &= \mathbf{(I - P_X)(P_W - P_X)}\\ &= \mathbf{(I - P_X) P_W} = \mathbf{P_W - P_X}\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{A\Sigma A\Sigma} &= \mathbf{P_W - P_X P_W - P_X}\\ &= \mathbf{P_W^2 + P_X^2 - P_W P_X - P_W P_X} = \mathbf{P_W + P_X - P_X - P_X}\\ &= \mathbf{P_W - P_X}\end{aligned}$$

Thus, $\mathbf{A\Sigma}$ is idempotent.
Rank($\mathbf{A}$) = Rank($\mathbf{P_W}$-$\mathbf{P_X}$) = rank($\mathbf{W}$) - Rank($\mathbf{X}$) = 5-3 = 2
$\lambda = \dfrac{\boldsymbol{\gamma}^\mathbf{T} \mathbf{W^T (I - P_X) P_W (I - P_X) W} \boldsymbol{\gamma}}{2\sigma^2} = \dfrac{\boldsymbol{\gamma}^\mathbf{T} \mathbf{W^T (P_W - P_X) W} \boldsymbol{\gamma}}{2\sigma^2}$
$\therefore \dfrac{\mathbf{y^T (I - P_X) P_W (I - P_X) y}}{\sigma^2} \sim \chi_1^2(\lambda)$

(c) Show that $F = \dfrac{c\mathbf{y}^T (I - P_X) P_W (I - P_X)\mathbf{y}}{\mathbf{y}^T (I - P_W)\mathbf{y}}$ has an $F-$ distribution for some constant c when model B is the correct model. Report a numerical value for c and degrees of freedom. (10 marks)

Ans.
Let $\mathbf{W_2} = \dfrac{\mathbf{y^T (I - P_W) y}}{\sigma^2}$
Here $\mathbf{A_2} = \dfrac{\mathbf{I - P_W}}{\sigma^2}$ and $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, thus $A\Sigma = I - P_W$ is clearly idempotent.
$\mathbf{A_2} \boldsymbol{\mu} = \dfrac{\mathbf{I - P_W}}{\sigma^2} \mathbf{W} \boldsymbol{\gamma} = 0$ since $\mathbf{P_W W} = \mathbf{W}$.
D.F = rank($\mathbf{I - P_W}$) = 9-5=4
Thus $\mathbf{W_2} \sim \chi_4^2$
Note that $\mathbf{W_1} = \dfrac{\mathbf{y^T (I - P_X) P_W (I - P_X) y}}{\sigma^2} \sim \chi_2^2(\lambda)$ and
$\mathbf{W_2} = \frac{\mathbf{y^T (I - P_W) y}}{\sigma^2} \sim \chi_4^2$ and

$$\begin{aligned}&\mathbf{(I - P_X) P_W (I - P_X)} (\sigma^2 \mathbf{I}) \mathbf{(I - P_W)}\\ &= \sigma^2 \mathbf{(P_W - P_X)(I - P_W)}\\ &= \sigma^2 \mathbf{(P_W - P_W^2 - P_X - P_X P_W)}\\ &= \sigma^2 \mathbf{(P_W - P_W - P_X - P_X)}\\ &= \mathbf{0}\end{aligned}$$

Thus, $\mathbf{y^T (I - P_W) y}$ and $\mathbf{y^T (I - P_X) P_W (I - P_X) y}$ are independent random variables.

Hence, $F = \dfrac{\mathbf{W_1}/2}{\mathbf{W_2}/4} = \frac{4\mathbf{W_1}}{2\mathbf{W_2}} \sim F_{2,4}(\lambda)$.
$c = 2$. The degrees of freedom are $(2, 4)$

Q2. Consider the model

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i (1)$$

where $\epsilon \sim NID(0, \sigma^2)$. This notation means that the random errors (and the observations) have normal distributions and satisfy the Gauss-Markov property.
Define

$$\mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \text{ and } \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \text{ and } \mathbf{P_X} = \mathbf{X}(\mathbf{X^T X})^{-1}\mathbf{X^T} \text{ and } \mathbf{P_1} = \mathbf{1}(\mathbf{1^T 1})^{-1}\mathbf{1^T}$$

(a) What are the distribution of the quadratic forms $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}$ and $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y}$. (10 marks)

Ans.

Note that under model 1, $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ where $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$

Let $\mathbf{A}_1 = \mathbf{P_1}$, $\mathbf{A}_2 = \mathbf{P_X} - \mathbf{P_1}$ and $\mathbf{A}_3 = \mathbf{I} - \mathbf{P_X}$. Then

- $\mathbf{A}_1, \mathbf{A}_2$ and $\mathbf{A}_3$ are all $n \times n$ symmetric matrices,
- $\mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 = \mathbf{I}$, and
- $\text{rank}(\mathbf{A}_1) + \text{rank}(\mathbf{A}_2) + \text{rank}(\mathbf{A}_3) = 1 + (2-1) + (n-2) = n$.

Then, by Cochran's Theorem, $\frac{1}{\sigma^2}\mathbf{Y^T A_i Y} \sim \chi^2_{r_i}(\frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T \mathbf{A}_i \mathbf{X}\boldsymbol{\beta})$
Since $\text{rank}(\mathbf{P_X} - \mathbf{P_1})$ = 2-1 = 1, then
$\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y} \sim \chi^2_1(\frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{P_X} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta})$
Since $\text{rank}(\mathbf{I} - \mathbf{P_X}) = n - 2$, then
$\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y} \sim \chi^2_{n-2}(\frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I} - \mathbf{P_X})\mathbf{X}\boldsymbol{\beta}) \sim \chi^2_{n-2}$ since $(\mathbf{I} - \mathbf{P_X})\mathbf{X}\boldsymbol{\beta} = 0$

(b) Derive the distribution of $F = \frac{(n-2)\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}}{\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y}}$. Report degrees of freedom and a formula for the noncentrality parameter. (10 marks)

Ans.

By Cochran's Theorem and part (a), $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}$ and $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y}$ are independent chi-square distributed with 1 and $n - 2$ df respevtively. Then,

$$\frac{\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}/2}{\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y}/(n-2)} = \frac{(n-2)\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}}{\mathbf{Y^T}(\mathbf{I} - \mathbf{P_X})\mathbf{Y}} \sim F_{1,n-2}(\lambda)$$

where $\lambda = \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{P_X} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta}$

(c) What is the null hypothesis associated with the F statistic in part (b)? Justify your answer by showing that the noncentrality parameter in part (b) is zero if and only if the null hypothesis is true. (10 marks)

Ans.

$$\begin{aligned}
\lambda &= \tfrac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{P_X} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta} \\
&= \tfrac{1}{\sigma^2}[(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta} - (\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I} - \mathbf{P_X})\mathbf{X}\boldsymbol{\beta}] \\
&= \tfrac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta} \text{ Since } \mathbf{I} - \mathbf{P_X})\mathbf{X} = \mathbf{0} \\
&= \tfrac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I} - \mathbf{P_1})(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta} \text{ since } (\mathbf{I} - \mathbf{P_1}) \text{ is idempotent} \\
&= \tfrac{1}{\sigma^2}[(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta}]^T(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta}
\end{aligned}$$

Note that

$$\begin{aligned}
(\mathbf{I} - \mathbf{P_1})\mathbf{X}\boldsymbol{\beta} &= \left[(\mathbf{I} - \mathbf{P_1})\mathbf{1}\ (\mathbf{I} - \mathbf{P_1}\mathbf{U})\right]\boldsymbol{\beta} \text{ where } \mathbf{U} = [X_1, \ldots, X_n]^T \\
&= \left[\mathbf{0}\ \mathbf{U} - \bar{\mathbf{U}}\mathbf{1}\right]\boldsymbol{\beta} \text{ where } \bar{\mathbf{U}} = \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \\
&= \beta_1(\mathbf{U} - \bar{\mathbf{U}}\mathbf{1})
\end{aligned}$$

Thus

$$\begin{aligned}
\lambda &= \tfrac{1}{\sigma^2}\beta_1^2(\mathbf{U} - \bar{\mathbf{U}}\mathbf{1})^T(\mathbf{U} - \bar{\mathbf{U}}\mathbf{1}) \\
&= \frac{\beta_1^2\sum_{i=1}^n(X_i - \bar{X})^2)}{\sigma^2}
\end{aligned}$$

Thus, $\lambda = 0$ if and only if $\beta_1 = 0$. Therefore, the null hypothesis is $H_0 : \beta_1 = 0$, and the non-centrality parameter λ is zero if and only if this null hypothesis is true.

(d) Suppose

$$\mathbf{Z} = \begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ \vdots & & \\ 1 & X_n & X_n^2 \end{bmatrix}.$$

Does $\frac{\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_1})\mathbf{Y}}{\mathbf{Y^T}(\mathbf{I} - \mathbf{P_Z})\mathbf{Y}/(n-3)}$, where $\mathbf{P_Z} = \mathbf{Z}(\mathbf{Z^T}\mathbf{Z})^{-1}\mathbf{Z}$, have an F-dsitribution when $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ is true? Explain. (10 marks)

Ans.

Let $\mathbf{A}_1 = \mathbf{P_1}$, $\mathbf{A}_2 = \mathbf{P_X} - \mathbf{P_1}$, $\mathbf{A}_3 = \mathbf{P_Z} - \mathbf{P_X}$ and $\mathbf{A}_4 = \mathbf{I} - \mathbf{P_Z}$. Then

- $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ and $\mathbf{A}_4$ are all $n \times n$ symmetric matrices,
- $\mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4 = \mathbf{I}$, and
- rank$(\mathbf{A}_1)$ + rank$(\mathbf{A}_2)$ + rank$(\mathbf{A}_3)$+ rank$(\mathbf{A}_4) = 1 + (2-1) + (3-2) + (n-3) = n$.

Then, by Cochran's Theorem, $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{I}-\mathbf{P_Z})\mathbf{Y} \sim \chi^2_{n-3}(\lambda_2)$where $\lambda_2 = \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I}-\mathbf{P_Z})\mathbf{X}\boldsymbol{\beta}$, and $\mathbf{Y^T A_1 Y}$, $\mathbf{Y^T A_2 Y}$, $\mathbf{Y^T A_3 Y}$ and $\mathbf{Y_4^A}$ are distributed independently.

$\frac{\mathbf{Y^T(P_X-P_1)Y}}{\mathbf{Y^T(I-P_Z)Y}/(n-3)}$ will have an F-distribution if $\mathbf{Y^T}(\mathbf{P_X}-\mathbf{P_1})\mathbf{Y}$ and $\mathbf{Y^T}(\mathbf{I}-\mathbf{P_Z})\mathbf{Y}$ are independent and $\lambda = 0$. $\mathbf{Y^T}(\mathbf{P_X}-\mathbf{P_1})\mathbf{Y}$ and $\mathbf{Y^T}(\mathbf{I}-\mathbf{P_Z})\mathbf{Y}$ are independent by Cochran's theorem. Now

$$\begin{aligned}
\lambda_2 &= \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I}-\mathbf{P_Z})\mathbf{X}\boldsymbol{\beta} \\
&= \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{I}-\mathbf{P_Z})\mathbf{P_X}\mathbf{X}\boldsymbol{\beta} \text{since } \mathbf{P_X X} = \mathbf{X} \\
&= \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{P_X}-\mathbf{P_Z P_X})\mathbf{X}\boldsymbol{\beta} \\
&= \frac{1}{\sigma^2}(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{P_X}-\mathbf{P_X})\mathbf{X}\boldsymbol{\beta} \text{ SInce } \mathbf{P_Z P_X} = \mathbf{P_X} \\
&= 0
\end{aligned}$$

Note that $\mathbf{P_Z P_X} = \mathbf{P_Z X(X^T X)^{-1} X^T} = \mathbf{X(X^T X)^{-1} X^T} = \mathbf{P_X}$ since $\mathbf{X} \in \mathcal{C}\mathbf{Z}$ and hance $\mathbf{P_Z X} = \mathbf{X}$

Q3. A food scientist performed the following experiment to study the effects of combining two different fats and three different surfactants on the specific volume of bread loaves. Four batches of dough were made for each of the six combinations of fat and surfactant. Ten loaves of bread were made from each batch of dough and the average volume of the ten loaves was recorded for each batch. Unfortunately, some of the yeast used to make some batches of dough was ineffective and data from the loaves made from those batches had to be removed from the analysis. Fortunately, all six combinations of the levels of fat and surfactant were observed at least once. The data (average volume of 10 loaves) are shown below.

	Surfactant		
	A	B	C
Fat 1	6.7	7.1	5.5
	4.3	5.9	6.4
	5.7	5.6	5.8
Fat 2	5.9	5.6	6.4
	7.4	6.8	5.1
	7.1	6.9	6.2
	7.0	7.2	6.3

Consider the model $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$ where $\epsilon_{ijk} \sim NID(0, \sigma^2)$ and Y_{ijk} denotes the average of the volumes of ten loaves of bread made from the k^{th} batch of dough using the i^{th} fat and the j^{th} surfactant. α_i is associated with the i-th level of fat, β_j is associated with the j-th level of surfactant, and γ_{ij} is an interaction parameter.

(a)Define $SSE = \sum_{i=1}^{2}\sum_{j=1}^{3}\sum_{k=1}^{4}(y_{ijk} - \bar{y}_{ij\bullet})^2$, where $\bar{y}_{ij\bullet} = \frac{1}{4}(y_{ij1} + y_{ij2} + y_{ij3} + y_{ij4})$. Show that $\frac{SSE}{\sigma^2}$ has a chi-squares distribution. States the degrees of fredom. (15 marks)

Ans.

Define $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where

$\mathbf{y} = [y_{111}, y_{112}, y_{113}, y_{114}, y_{121}, y_{122}, y_{123}, y_{124}, y_{131}, y_{132}, y_{133}, y_{134}, y_{211}, y_{212}, y_{213}, y_{214}, y_{221}, y_{222},$

$\boldsymbol{\beta} = [\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{13}]^T$ and $\mathbf{X}$ is the corresponding model matrix. Then

$\frac{SSE}{\sigma^2} = \mathbf{y^T}\frac{\mathbf{(I-P_X)}}{\sigma^2}\mathbf{y}$ where $\mathbf{P_X} = \mathbf{X}(\mathbf{X^TX})^{-}\mathbf{X^T}$.

Here $\mathbf{A} = y^T\frac{\mathbf{(I-P_X)}}{\sigma^2}$ and $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, then

$\mathbf{A\Sigma} = \mathbf{I} - \mathbf{P_X}$ clearly is idempotent.

$\mathbf{A}\boldsymbol{\mu} = \frac{1}{\sigma^2}(\mathbf{I} - \mathbf{P_X})\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$.

$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{I} - \mathbf{P_X}) = n - \text{rank}(\mathbf{X}) = 24\text{-}6 = 18$

$\therefore \frac{SSE}{\sigma^2} \sim \chi^2_{18}$.

(b)Consider the estimator

$$\hat{C} = \bar{y}_{1\bullet\bullet} - \bar{y}_{2\bullet\bullet},$$

where

$$\bar{y}_{i\bullet\bullet} = \frac{1}{8}\sum_{j=1}^{2}\sum_{k=1}^{4} y_{ijk}.$$

Show that

$$F = \frac{m(\hat{C})^2}{SSE}$$

has an F-distribution for some constant m. Report the value of m and the degrees of freedom for the F-distribution. (15 marks)

Ans.

Let $\mathbf{a^T} = \frac{1}{12}[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1,]$,

we have $\widehat{\mathbf{C}} = \mathbf{a^Ty}$ and $\widehat{\mathbf{C}}^2 = \mathbf{y^Taa^Ty}$

Take $\mathbf{A} = \frac{\mathbf{aa^Ta^Ta}}{\sigma^2} = \frac{6\mathbf{aa}^T}{\sigma^2}$ and $\mathbf{\Sigma} = \sigma^2\mathbf{I}$.

Then $\mathbf{A\Sigma} = 6\mathbf{aa^T}$.

$\mathbf{A\Sigma A\Sigma} = (6\mathbf{aa^T})(6\mathbf{aa^T}) = 36\mathbf{aa^Taa^T} = 36\mathbf{a}(\frac{1}{6})\mathbf{a^T} = 6\mathbf{aa^T}$

$\therefore$ $\mathbf{A\Sigma}$ is idempotent.

$\mathbf{a^T}\boldsymbol{\mu} = \mathbf{a^TX}\boldsymbol{\beta} = \alpha_1 - \alpha_2 + \bar{\gamma}_{1\bullet} + \gamma_{2\bullet}$, thus

$\lambda = \boldsymbol{\mu}^\mathbf{T}\mathbf{aa^T}\boldsymbol{\mu} = \frac{1}{\sigma^2}(\alpha_1 - \alpha_2 + \bar{\gamma}_{1\bullet} + \gamma_{2\bullet})^2$

$\text{Rank}(\mathbf{A}) = \text{Rank}(\mathbf{aa^T}) = 1$

$\therefore \frac{6(\widehat{\mathbf{C}})^2}{\sigma^2} = \frac{6\mathbf{y^Taa^Ty}}{\sigma^2} \sim \chi^2_1(\lambda)$

To show that $\widehat{\mathbf{C}}^2$ is independent of SSE. Let $\mathbf{A_1} = \mathbf{I} - \mathbf{P_X}$, $\mathbf{A_2} = \mathbf{aa^T}$.

$$\begin{aligned}\mathbf{A_1\Sigma A_2} &= (\mathbf{I} - \mathbf{P_X})\sigma^2\mathbf{I}(\mathbf{aa^T}) \\ &= \sigma^2[\mathbf{aa^T} - \mathbf{P_Xaa^T}] \\ &= \sigma^2[\mathbf{aa^T} - \mathbf{aa^T}] \text{ since } \mathbf{a} \in \mathcal{C}(\mathbf{X}) \\ &= \mathbf{0}\end{aligned}$$

Thus $\widehat{\mathbf{C}}^2$ and SSE are independent. Consequently,

$$F = \frac{\frac{6(\widehat{\mathbf{C}})^2}{\sigma^2}/1}{\frac{SSE}{\sigma^2}/18}$$
$$= \frac{108(\widehat{\mathbf{C}})^2}{SSE} \sim F_{1,18}(\lambda)$$

where $\lambda = \frac{1}{\sigma^2}(\alpha_1 - \alpha_2 + \bar{\gamma}_{1\bullet} + \gamma_{2\bullet})$